Counting Markov Types, Balanced Matrices, and Eulerian Graphs[*]

February 15, 2012

Philippe Jacquet
Bell Labs
Alcatel-Lucent
91620 Nozay
France
philippe.jacquet@inria.fr

Charles Knessl[†]
Dept. Math. Stat. & Compt. Sci.
University of Illinois at Chicago
Chicago, Illinois 60607-7045
U.S.A.
knessl@uic.edu
.

Wojciech Szpankowski[‡]
Dept. Computer Science
Purdue University
W. Lafayette, IN 47907
U.S.A.
spa@cs.purdue.edu

Abstract

The method of types is one of the most popular techniques in information theory and combinatorics. Two sequences of equal length have the same type if they have identical empirical distributions. In this paper, we focus on Markov types, that is, sequences generated by a Markov source (of order one). We note that sequences having the same Markov type share the same so called *balanced frequency matrix* that counts the number of distinct pairs of symbols. We enumerate the number of Markov types for sequences of length n over an alphabet of size m. This turns out to be asymptotically equivalent to estimating the number of the balanced frequency matrices, the number of integer solutions of a system of *linear Diophantine equations*, and the number of connected Eulerian multigraphs. For fixed m we prove that the number of Markov types is asymptotically equal to

$$d(m)\frac{n^{m^2-m}}{(m^2-m)!},$$

where we give an integral representation for $d(m)$. For $m \to \infty$ we conclude that asymptotically the number of types is equivalent to

$$\frac{\sqrt{2}m^{3m/2}e^{m^2}}{m^{2m^2}2^m\pi^{m/2}}n^{m^2-m}$$

provided that $m = o(n^{1/4})$. These findings are derived by analytical techniques ranging from analytic combinatorics, to multidimensional generating functions, to the saddle point method.

Index Terms: Markov types, Eulerian graphs, balance frequency matrices, linear Diophantine equations, multidimensional generating functions, saddle point method.

1 Introduction

The *method of types* is one of the most popular and useful techniques in information theory and combinatorics. Two sequences of equal length are of the same type if they have identical empirical distributions. Furthermore, sequences of the same type are assigned the same probability by all distributions in a given class. The essence of the method of types was known for some time in probability and statistical physics. But only in the 1970's Csiszár and his group developed a general method and made it a basic tool of information theory of discrete memoryless systems [5]; see also [4, 6, 9, 14, 15, 17, 23, 24]. In passing we observe that the number of types are often needed for minimax redundancy evaluations [1, 2, 14, 16, 18, 22].

There are basically two equivalent ways to define types, one combinatorial [4, 6] and one probabilistic [6, 17]. Throughout the paper we use a combinatorial approach. In general, sequences that have the same *empirical distribution* belong to the same type class. More precisely, let $\mathcal{A} = \{1, 2, \ldots, m\}$ be an m-ary alphabet. For a given sequence of length n, $x^n = x_1 \ldots x_n \in \mathcal{A}^n$, the *type class* is defined as

$$\mathcal{T}_n(x^n) = \{y^n : \ P_{x^n} = P_{y^n}\}$$

for all empirical distributions P_{x^n} in a given model class. Clearly, $\bigcup_{x^n} \mathcal{T}_n(x^n) = \mathcal{A}^n$, and $|\mathcal{T}_n(x^n)|$ counts the number of sequences of the same type as x^n. We denote the set of all types as $\mathcal{Q}_n(m)$. We aim at deriving an asymptotic expression for the number of (first order) Markov types $|\mathcal{Q}_n(m)|$ (defined more precisely below) for $n \to \infty$, considering both fixed and large m.

Here is another equivalent combinatorial characterization of types based on symbolic calculus. We start with an example illustrating types for memoryless sources. Consider alphabet $\mathcal{A} = \{1, 2, \ldots, m\}$ and m "formal parameters" $p_1, p_2, \ldots, p_m$. From a combinatorial (symbolic) point of view, $p_1, \ldots, p_m$ can be viewed as "formal indeterminate". Then for a given $x^n \in \mathcal{A}^n$ define formally a "pattern"

$$P(x^n) := p_1^{k_1} \cdots p_m^{k_m}, \qquad k_1 + k_2 + \cdots + k_m = n,$$

where k_i is the number of times symbol $i \in \mathcal{A}$ occurs in x^n. Here, a type is fully characterized by the *vector count* $\mathbf{k} = (k_1, \ldots, k_m)$ such that $k_1 + k_2 \cdots + k_m = n$. Thus, for a given x^n, or a given count $\mathbf{k}(x^n) = (k_1(x^n), \ldots, k_m(x^n))$, the type class $\mathcal{T}(x^n) = \mathcal{T}(\mathbf{k})$ is fully described by the "formal pattern" $p_1^{k_1} \cdots p_m^{k_m}$. It consists of all sequences x^n of the same count or pattern. In particular,

$$|\mathcal{T}(\mathbf{k})| = \binom{n}{k_1, \ldots, k_m}, \quad |\mathcal{Q}_n(m)| = \binom{n+m-1}{m-1}.$$

Observe that if we interpret $p_1, \ldots, p_m$ as probabilities of individual symbols in $\mathcal{A}$ (i.e., $p_1 + \cdots + p_m = 1$), then $P(x^n)$ becomes the probability of x^n, and sequences of the same type are assigned the same probability in the class of memoryless sources.

Let us now consider Markovian types using our symbolic approach. We first define a matrix of formal indeterminate $\mathbf{P} = \{p_{ij}\}_{i,j \in \mathcal{A}}$. For a given x^n define formally

$$P(x^n) := \prod_{i,j \in \mathcal{A}} p_{ij}^{k_{ij}}, \tag{1}$$

2

where k_{ij} is the number of the pairs $(ij) \in \mathcal{A}^2$ in x^n. For example, $P(01011) = p_{01}^2 p_{10} p_{11}$. The frequency matrix $\mathbf{k}$ is an integer matrix satisfying two important properties:

$$\sum_{i,j \in \mathcal{A}} k_{ij} = n - 1, \tag{2}$$

and additionally for any $i \in \mathcal{A}$ [14, 24]

$$\sum_{j=1}^{m} k_{ij} = \sum_{j=1}^{m} k_{ji} + \delta(x_1 = i) - \delta(x_n = i), \quad \forall i \in \mathcal{A}, \tag{3}$$

where $\delta(A) = 1$ when A is true and zero otherwise. The last property is called the *flow conservation property* and is a consequence of the fact that the number of pairs starting with symbols $i \in \mathcal{A}$ must be equal to the number of pairs ending with symbol $i \in \mathcal{A}$ with the possible exception of the first and last pairs. To avoid this exception, we first consider *cyclic* strings in which the first element x_1 follows the last x_n (cf. also [10]). Thus, we consider integer matrices $\mathbf{k} = [k_{ij}]$ satisfying the following two properties

$$\sum_{i,j \in \mathcal{A}} k_{ij} = n, \tag{4}$$

$$\sum_{j=1}^{m} k_{ij} = \sum_{j=1}^{m} k_{ji}, \quad \forall\, i \in \mathcal{A}. \tag{5}$$

Such integer matrices $\mathbf{k}$ will be called *balanced frequency matrices* or simply balanced matrices. We shall call (5) the "conservation law" equation or simply the *balanced boundary condition* (BBC). We denote by $\mathcal{F}_n(m)$ the set of nonnegative integer solutions of (4) and (5).

Now we are ready to define cyclic Markov types. Two cyclic sequences have the same (cyclic) Markov type if they have the same formal pattern $P(x^n)$ defined in (1). We denote by $\mathcal{P}_n(m)$ the set of cyclic Markov types and enumerate them by comparing them to the cardinality of $\mathcal{F}_n(m)$.

Let us now briefly address a probabilistic interpretation of Markov types. If p_{ij} represent transition probabilities and the initial condition is fixed, then $P(x^n)$ is the probability of x^n. In the cyclic setting described above, the initial condition is a cyclic one; we address non-cyclic sequences (i.e., arbitrary initial condition) and the corresponding types in Section 2.3. Note that two sequences x^n and y^n of the same Markov type (i.e., the same frequency matrix, and the same initial condition) are assigned the same probability by all Markov sources (of order one). Hereafter (with the exception of Section 2.3) we only deal with cyclic Markov types $\mathcal{P}_n(m)$ – that we shall simply call Markov types – since they allow us for an elegant presentation of combinatorial and analytic methodologies that relate Markov types to a few other interesting combinatorial objects, and may possibly lead to new results in other studies such as types of Markov fields.

A (cyclic) Markov type is characterized by its balanced frequency matrix $\mathbf{k}$, but not every $\mathbf{k}$ leads to a sequence having a Markov type since there may not be an underlying legitimate sequence x^n (see Example 2 below). To better see this, we present another characterization of Markov types. Let us define a directed multigraph $G = (V, E)$ with the set of vertices $V = \mathcal{A}$ and k_{ij} edges between vertices $i, j \in \mathcal{A}$. For $\mathcal{A} = \{0, 1\}$ such a graph

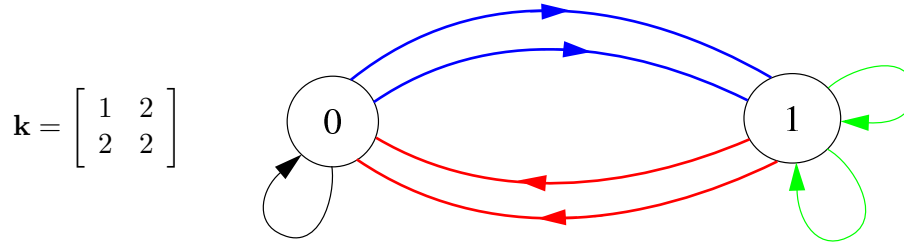$$\mathbf{k} = \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix}$$

Figure 1: A frequency matrix and its corresponding Eulerian graph.

is shown in Figure 1. Then, as already observed in [3, 13, 14], the number of sequences of a given type $\mathbf{k}$, i.e., $|\mathcal{T}(\mathbf{k})|$, is equal to the number of Eulerian cycles in G. On the other hand, the number of types $|\mathcal{P}_n(m)|$ coincides with the number of Eulerian digraphs $G = (V, E)$ such that $V \subseteq \mathcal{A}$ and $|E| = n$ (here $V \subseteq \mathcal{A}$ since there may be sequences composed of only *some* symbols of the alphabet). In this paper, we enumerate the number of Markov types $|\mathcal{P}_n(m)|$, that is, the number of Eulerian digraphs G on $V \subseteq \mathcal{A}$. In Section 2.1 we show that the number of types, $|\mathcal{P}_n(m)|$ is asymptotically equivalent to the number of integer solutions $|\mathcal{F}_n(m)|$ of (4) and (5).

Example 1. Let's first consider a binary Markov source. A balanced frequency matrix is of the following form

$$\mathbf{k} = \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix}$$

where the nonnegative integers k_{ij} satisfy

$$\begin{aligned} k_{11} + k_{12} + k_{21} + k_{22} &= n, \\ k_{12} &= k_{21}. \end{aligned}$$

From the above we conclude that

$$k_{11} + 2k_{12} + k_{22} = n. \tag{6}$$

The number of nonnegative integer solutions of (6) is obviously

$$\begin{aligned} |\mathcal{F}_n(2)| &= \sum_{k_{12}=0}^{\lfloor \frac{n}{2} \rfloor} (n - 2k_{12} + 1) \\ &= \left(\left\lfloor \frac{n}{2} \right\rfloor + 1 \right) \left(n - \left\lfloor \frac{n}{2} \right\rfloor + 1 \right) = \frac{n^2}{4} + O(n) \end{aligned} \tag{7}$$

which we shall show is asymptotically the same as the number of types. We note that the number, $|\mathcal{T}(\mathbf{k})|$, of circular sequences of length n is equivalent to the number of Eulerian cycles of the corresponding graph G, as shown in Figure 1 (see [14] for a detailed discussion).

Example 2. Let's now look at the $m = 3$ case. The balanced frequency matrix has nine elements $\{k_{ij}\}_{i,j \in \{1,2,3\}}$, and they satisfy

$$k_{11} + k_{12} + k_{13} + k_{21} + k_{22} + k_{23} + k_{31} + k_{32} + k_{33} = n$$

4

$$\begin{aligned}
k_{12} + k_{13} &= k_{21} + k_{31} \\
k_{12} + k_{32} &= k_{21} + k_{23} \\
k_{13} + k_{23} &= k_{31} + k_{32}.
\end{aligned}$$

How many nonnegative integer solutions does the above system of linear equations have? The answer is not quite obvious, but we shall show that it is asymptotically $\frac{n^6}{12 \cdot 6!}$. However, we observe that not every solution of the above system of equations leads to a legitimate sequence x^n. Consider the following example with $n = 5$

$$\begin{aligned}
k_{11} &= 2, \quad k_{12} = k_{13} = 0, \\
k_{21} &= 0, \quad k_{22} = 0, \quad k_{23} = 1, \\
k_{31} &= 0, \quad k_{32} = k_{33} = 1
\end{aligned}$$

There is no sequence of length five that can satisfy the above conditions. For example, if $\mathcal{A} = \{a, b, c\}$ such a cyclic sequence would have to consist of five elements and contain the substrings bc, cb, cc, and two $aa's$ or one aaa, which is impossible. Observe that the underlying multigraph consists of two connected components that do not communicate. Nevertheless, in the next section we prove that the number of integer solutions of the above system of equations coincides *asymptotically* with the number of types $|\mathcal{P}_n(3)|$. ∎

Our goal is to enumerate the number of Markov types, that is, to find the cardinality of $|\mathcal{P}_n(m)|$ (and $|\mathcal{Q}_n(m)|$) by first enumerating $|\mathcal{F}_n(m)|$, the number of nonnegative integer solutions to the system of linear equations (4)-(5). Such an enumeration, for a general class of systems of homogeneous Diophantine equations, was investigated in Chap. 4.6 of Stanley's book [19] (cf. also [12]). Stanley developed a general theory to construct the associated generating function. However, ultimately only the denominator of this generating function is given in a semi-explicit form in [19], which suffices to derive the growth rate of the number of integer solutions.

In this paper, we propose an approach based on previous work of Jacquet and Szpankowski [14], where analytic techniques such as multidimensional generating functions and the saddle point method were used. This allows us to derive precise asymptotic results. In particular, for fixed m we establish in Theorem 1 that the number of Markov types is asymptotically equal to

$$|\mathcal{P}_n(m)| \sim d(m) \frac{n^{m^2 - m}}{(m^2 - m)!}, \quad n \to \infty,$$

and we shall give an integral representation for $d(m)$. We also show that the number of types $|\mathcal{Q}_n(m)|$ over non-cyclic Markov sequences (i.e., with arbitrary initial condition) is asymptotically equivalent to $|\mathcal{P}_n(m)|$ but larger by the factor $(m^2 - m + 1)$ (see Corollary 1). For large $m \to \infty$ with $m^4 = o(n)$ we determine $d(m)$ and find that asymptotically the number of types is

$$|\mathcal{P}_n(m)| \sim \frac{\sqrt{2} m^{3m/2} e^{m^2}}{m^{2m^2} 2^m \pi^{m/2}} n^{m^2 - m}.$$

Moreover, our techniques also allow us to derive asymptotics for m to be as large as $m^2 = O(n)$. Finally, we point out that our technique easily generalizes to Markov types of order

r. It can be proved (e.g., see [11]) that

$$|\mathcal{P}_n^r(m)| \sim d_r(m) \frac{n^{m^r(m-1)}}{[m^r(m-1)]!} \qquad (8)$$

where $d_r(m)$ is a complicated constant. For example in [11] it was shown that $d_2(2) = 1/12$.

Markov types were studied in a series of papers; see [14, 15, 23, 24]. However, the existing literature mostly concentrates on finding the number of strings of a given Markov type class, that is, $|\mathcal{T}(x^n)|$; an exception is the work of Martin et al. [15]. In particular, Whittle [24] already in 1955 computed $|\mathcal{T}(x^n)|$ for Markov chains of order one. Regarding the number of types, it was known for some time [4, 5, 6] that they grow polynomially in n, but only in [23] Weinberger et al. mentioned (without proof) that $|\mathcal{P}_n| = \Theta(n^{m^2-m})$. This estimate was recently rigorously proved by Martin et al. in [15] for tree sources (that include Markov sources) for fixed m. However, the constant in the asymptotic estimate was never identified. We accomplish it here, as well as present asymptotic results for large m.

The paper is organized as follows. In Section 2 we formulate precisely our problem, establish some asymptotic equivalences, and present our main asymptotic results for fixed or large m. The proofs are given in Section 3.

2 Main Results

In this section we present our main results and some of their consequences. We first derive some asymptotic equivalences. In the introduction, we defined the set $\mathcal{F}_n(m)$ of all solutions of the balance equations (4) and (5). We also noticed that the number of (cyclic) types $|\mathcal{P}_n(m)|$ is related to the number of Eulerian (connected) directed multigraphs (we shall call them Eulerian digraphs, for short). In Lemma 1 of this section we prove that $|\mathcal{F}_n(m)|$ and $|\mathcal{P}_n(m)|$ are asymptotically equivalent. Then we develop a method to evaluate asymptotically $|\mathcal{F}_n(m)|$ which leads to our main result, presented in Theorem 1. We then briefly discuss non-cyclic Markov types $\mathcal{Q}_n(m)$ summarizing our findings in Corollary 1.

2.1 Some Asymptotic Equivalences

Recall that $\mathcal{P}_n(m)$ stands for the set of (cyclic) Markov types of length n over an alphabet $\mathcal{A}$ of size m. As already observed, it is equal to the set of all connected Eulerian di-graphs $G = (V(G), E(G))$ such that $V(G) \subseteq \mathcal{A}$ and $|E(G)| = n$. The point we emphasize is that G may be defined over a *subset* of $\mathcal{A}$, as shown in the first example in Figure 2 (i.e., there may be some isolated vertices). We denote by $\mathcal{E}_n(m)$ the set of connected Eulerian digraphs on $\mathcal{A}$; the middle of Figure 2 shows an example of a graph in this set. Finally, the set $\mathcal{F}_n(m)$ can be viewed as the set of digraphs G with $V(G) = \mathcal{A}$, $|E(G)| = n$ and satisfying the flow conversation property (in-degree equals out-degree). We call such graphs conservative digraphs. Observe that a graph in $\mathcal{F}_n(m)$ may consist of several connected (not communicating) Eulerian digraphs, as shown in the third example in Figure 2.

There is a simple relation between $|\mathcal{E}_n(m)|$ and $|\mathcal{P}_n(m)|$. Indeed,

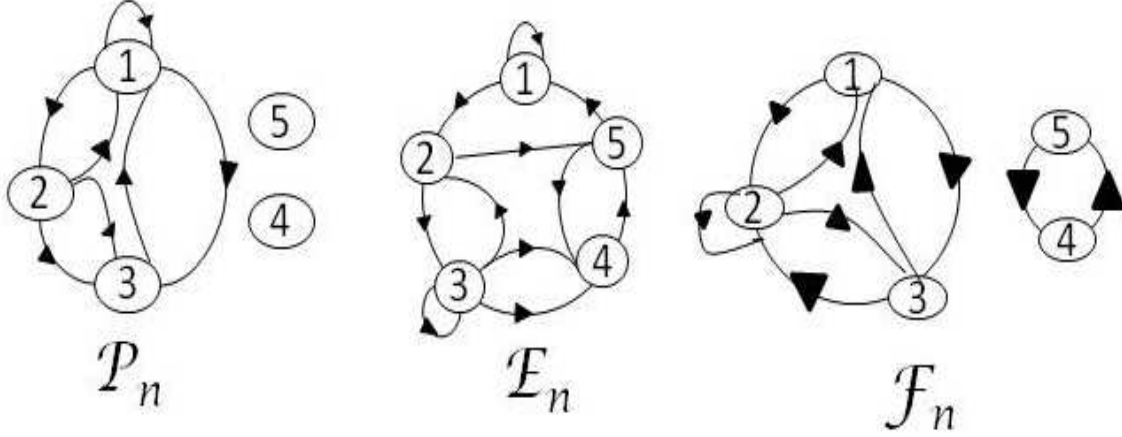$$|\mathcal{P}_n(m)| = \sum_k \binom{m}{k} |\mathcal{E}_n(k)| \qquad (9)$$

Figure 2: Examples of graphs belonging to $\mathcal{P}_7(5)$, $\mathcal{E}_{11}(5)$ and $\mathcal{F}_9(5)$ sets.

since there are $\binom{m}{k}$ ways to choose $m - k$ isolated vertices in $\mathcal{P}_n(m)$. This implies that

$$|\mathcal{P}_n(m)| \geq |\mathcal{E}_n(m)|.$$

Indeed, it is certainly true since every Eulerian graph on $\mathcal{A}$ with $k \leq m$ (by definition it must be strongly connected and satisfy the flow-conservation law) belongs to $\mathcal{P}_n(m)$. In fact, by the same reasoning we can expand the above inequality to obtain for any n and m

$$|\mathcal{F}_n(m)| \geq |\mathcal{P}_n(m)| \geq |\mathcal{E}_n(m)|. \tag{10}$$

We now find an asymptotic relation between the number of digraphs $|\mathcal{P}_n(m)|$ and the number of solutions $|\mathcal{F}_n(m)|$ of the flow conservation equations (4)-(5), that is, the number of conservative digraphs. As a direct consequence of our definition, a conservative digraph may have several connected components. Each connected component is either a connected Eulerian digraph or an isolated node without an edge. This leads to

$$|\mathcal{F}_n(m)| = |\mathcal{E}_n(m)| + \sum_{i=2}^{m} \sum_{\mathcal{A}=\mathcal{A}_1\cup\cdots\cup\mathcal{A}_i} \sum_{n_1+\cdots+n_i=n} \prod_{j=1}^{i} |\mathcal{E}_{n_j}(\mathcal{A}_j)| \tag{11}$$

where the sum is over all (unordered) set partitions $\mathcal{A} = \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_i$ into $i \geq 2$ (nonempty) parts with n_j edges in each di-subgraph $\mathcal{E}_{n_j}(\mathcal{A}_j)$ over $\mathcal{A}_j$ vertices. Observe that every set partition $\mathcal{A} = \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_i$ with $|\mathcal{A}_j| = m_j > 0$ is a partition of $\mathcal{A}$ into i distinguished subsets of cardinality m_j. Notice that there are $\binom{m}{m_1...m_i}$ ways of dividing m into i subsets of size m_i when permuting the subsets does not lead to a new distinct permutation. For example, for $\mathcal{A} = \{1, 2, 3, 4\}$ and $m_1 = m_2 = 2$, we have the following partition of $\mathcal{A}$:

$$\{1, 2\}, \{3, 4\}; \ \{1, 3\}, \{2, 4\}; \ \{2, 3\}, \{1, 4\}.$$

But there are three additional partitions included in the count $\binom{4}{2,2} = 6$, namely

$$\{3, 4\}, \{1, 2\}; \ \{2, 4\}, \{1, 3\}; \ \{1, 4\}, \{2, 3\}$$

7

that follow from the above by permuting equal subsets. In general, permuting two subsets of equal size does not lead to a new partition, as seen above. As a consequence of this we can write (11) as

$$|\mathcal{F}_n(m)| \le |\mathcal{E}_n(m)| + \sum_{i=2}^{m} \sum_{m_1+\cdots+m_i=m} \sum_{n_1+\cdots+n_i=n} \binom{m}{m_1 \ldots m_i} \prod_{j=1}^{i} |\mathcal{E}_{n_j}(m_j)|. \tag{12}$$

Furthermore, since by (10) $|\mathcal{E}_n(m)| \le |\mathcal{F}_n(m)|$ for all $n, m \ge 0$ we finally arrive at

$$|\mathcal{F}_n(m)| \le |\mathcal{E}_n(m)| + \sum_{i=2}^{m} \sum_{m_1+\cdots+m_i=m} \sum_{n_1+\cdots+n_i=n} \binom{m}{m_1 \ldots m_i} \prod_{j=1}^{i} |\mathcal{F}_{n_j}(m_j)|. \tag{13}$$

Now, we are in the position to formulate our asymptotic equivalence result.

Lemma 1 *The following holds for all $m \ge 2$ and $n \to \infty$*

$$||\mathcal{F}_n(m)| - |\mathcal{P}_n(m)|| = O(2^m m^3 n^{m^2-3m+3}). \tag{16}$$

Proof. By (10) we know that $|\mathcal{P}_n(m)| \le |\mathcal{F}_n(m)|$, so we need only to prove that

$$|\mathcal{F}_n(m)| = |\mathcal{P}_n(m)| + O(2^m m^3 n^{m^2-3m+3}). \tag{17}$$

Our starting point is (13) where we denote the sum on the right-hand side by $A(m,n)$, that is,

$$A(m,n) = \sum_{i=2}^{m} \sum_{m_1+\cdots+m_i=m} \sum_{n_1+\cdots+n_i=n} \binom{m}{m_1 \ldots m_i} \prod_{j=1}^{i} |\mathcal{F}_{n_j}(m_j)|.$$

We need to prove that $A(m,n) = O(2^m m^3 n^{m^2-3m+3})$. In Theorem 1 below we shall prove that $|\mathcal{F}_n(m)| = O(n^{m^2-m})$, thus

$$\prod_{j=1}^{i} |\mathcal{F}_{n_j}(m_j)| = O(n^{m_1^2+\cdots+m_i^2-m}).$$

For $m_j \ge 1$ we have

$$\sum_{j=1}^{i} m_j^2 = m^2 - \sum_{j=1}^{i} m_j(m - m_j)$$

$$\le m^2 - \sum_{j=1}^{i}(m-1) = m^2 - i(m-1) \,,$$

and for some constant C (that may vary from line to line)

$$A(m,n) \le C \sum_{i=2}^{m} \sum_{m_1+\cdots+m_i=m} \sum_{n_1+\cdots+n_i=n} \binom{m}{m_1 \ldots m_i} n^{m^2-i(m-1)-m}$$

$$= Cn^{m^2-m} \sum_{i=2}^{m} n^{-i(m-1)} \sum_{n_1+\cdots+n_i=n} \sum_{m_1+\cdots+m_i=m} \binom{m}{m_1 \ldots m_i}$$

8

$$
\begin{aligned}
&= &Cn^{m^2-m}\sum_{i=2}^{m}n^{-i(m-1)}\sum_{n_1+\cdots+n_i=n}i^m\\
&\leq &Cn^{m^2-m}\sum_{i=2}^{m}n^{-i(m-1)}i^m n^{i-1}\\
&\leq &Cn^{m^2-m-1}\sum_{i=2}^{m}n^{-i(m-2)}i^m\\
&\leq &Cn^{m^2-m-1}m^2\sum_{i=2}^{m}\left(\frac{i}{n^i}\right)^{m-2}\\
&\leq &C2^m m^3 n^{m^2-3m+3}
\end{aligned}
$$

where the last line follows from the fact that $i/n^i \leq 2/n^2$ for $i \geq 2$. This completes the proof.

∎

Remark. Actually, we can find the exact relation between $|\mathcal{F}_n(m)|$ and $|\mathcal{E}_n(m)|$. Let

$$
P(z,u) = \sum_{m,n}|\mathcal{P}_n(m)|\frac{u^m z^n}{m!}, \quad E(z,u) = \sum_{m,n}|\mathcal{E}_n(m)|\frac{u^m z^n}{m!}.
$$

From (9) we find that

$$
P(z,u) = e^u E(z,u).
$$

Furthermore, define the bivariate generating function

$$
F(z,u) = \sum_{m\geq 1,n}|\mathcal{F}_n(m)|\frac{u^m z^n}{m!}.
$$

Recall that $\mathcal{F}_n(m)$ consists of Eulerian digraphs while $\mathcal{E}_n(m)$ is the set of *connected* Eulerian digraphs. The so called *exponential formula* [12] (page 118) relates $F(z,u)$ and $E(z,u)$, namely

$$
F(z,u) = \exp(E(z,u)) - 1
$$

where -1 represents the fact that the generating function of $|\mathcal{F}_n(m)|$ starts from $m \geq 1$. Thus

$$
E(z,u) = \log(1 + F(z,u))
$$

which translates into

$$
|\mathcal{E}_n(m)| = \sum_{i\geq 1}\frac{(-1)^{i+1}}{i}\sum_{m_1+\cdots+m_i=m}\binom{m}{m_1\cdots m_i}\sum_{n_1+\cdots+n_i=n}\prod_{j=1}^{i}|\mathcal{F}_{n_j}(m_j)|.
$$

Since

$$
F(z,u) = \sum_{i=1}^{\infty}\frac{1}{i!}E^i(z,u)
$$

we also find

$$
|\mathcal{F}_n(m)| = |\mathcal{E}_n(m)| + \sum_{i=2}^{m}\frac{1}{i!}\sum_{m_1+\cdots+m_i=m}\binom{m}{m_1\cdots m_i}\sum_{n_1+\cdots n_i=n}\prod_{j=1}^{i}|\mathcal{E}_{n_j}(m_j)|.
$$

Clearly, we could also use these expressions to establish our Lemma 1.

2.2 Counting Cyclic Markov Types

In Lemma 1 we proved that as $n \to \infty$

$$|\mathcal{P}_n(m)| \sim |\mathcal{F}_n(m)|.$$

Therefore, in the sequel we concentrate on estimating $|\mathcal{F}_n(m)|$.

We first make some general observations about generating functions over matrices, and summarize some results obtained in [14]. In general, let $g_{\mathbf{k}}$ be a sequence of scalars indexed by matrices $\mathbf{k}$ and define the generating function

$$G(\mathbf{z}) = \sum_{\mathbf{k}} g_{\mathbf{k}} \mathbf{z}^{\mathbf{k}}$$

where the summation is over all integer matrices and $\mathbf{z} = \{z_{ij}\}_{i,j \in \mathcal{A}}$ is an $m \times m$ matrix that we often denote simply as $\mathbf{z} = [z_{ij}]$ (assuming the indices i and j run from 1 to m). Here $\mathbf{z}^{\mathbf{k}} = \prod_{i,j} z_{ij}^{k_{ij}}$ where k_{ij} is the entry in row i column j in the matrix $\mathbf{k}$. We denote by

$$G^*(\mathbf{z}) = \sum_{\mathbf{k} \in \mathcal{F}} g_{\mathbf{k}} \mathbf{z}^{\mathbf{k}} = \sum_{n \geq 0} \sum_{\mathbf{k} \in \mathcal{F}_n(m)} g_{\mathbf{k}} \mathbf{z}^{\mathbf{k}}$$

the generating function of $g_{\mathbf{k}}$ over matrices $\mathbf{k} \in \mathcal{F}_n(m)$ satisfying the balance equations (4) and (5). The following useful lemma relates $G(\mathbf{z})$ and $G^*(\mathbf{z})$ is proved in [14] but for completeness we repeat it here. Let $[z_{ij} \frac{x_i}{x_j}]$ be the matrix $\Delta^{-1}(x)\mathbf{z}\Delta(x)$ where $\Delta(x) = \mathrm{diag}(x_1, \ldots, x_m)$ is a diagonal matrix with elements $x_1, \ldots, x_m$, that is, the element z_{ij} in $\mathbf{z}$ is replaced by $z_{ij} x_i / x_j$.

Lemma 2 *Let $G(\mathbf{z}) = \sum_{\mathbf{k}} g_{\mathbf{k}} \mathbf{z}^{\mathbf{k}}$ be the generating function of a complex matrix $\mathbf{z}$. Then*

$$G^*(\mathbf{z}) := \sum_{n \geq 0} \sum_{\mathbf{k} \in \mathcal{F}_n} g_{\mathbf{k}} \mathbf{z}^{\mathbf{k}} = \left(\frac{1}{2\mathbf{i}\pi}\right)^m \oint \frac{dx_1}{x_1} \cdots \oint \frac{dx_m}{x_m} G([z_{ij} \frac{x_j}{x_i}]) \qquad (18)$$

$$= \left[x_1^0 \cdots x_m^0\right] g([z_{ij} \frac{x_j}{x_i}])$$

with the convention that the ij-th coefficient of $[z_{ij} \frac{x_j}{x_i}]$ is $z_{ij} \frac{x_j}{x_i}$, and $\mathbf{i} = \sqrt{-1}$. In other words, $[z_{ij} \frac{x_j}{x_i}] = \Delta^{-1}(x)\mathbf{z}\Delta(x)$ where $\Delta(x) = \mathrm{diag}(x_1, \ldots, x_m)$. By the change of variables $x_i = \exp(\mathbf{i}\theta_i)$ we also have

$$G^*(\mathbf{z}) = \frac{1}{(2\pi)^m} \int_{-\pi}^{\pi} d\theta_1 \cdots \int_{-\pi}^{\pi} d\theta_m G([z_{ij} \exp((\theta_j - \theta_i)\mathbf{i})])$$

where $[z_{ij} \exp(\theta_j - \theta_i)] = \exp(-\Delta(\theta))\mathbf{z}\exp(\Delta(\theta))$.

Proof. Observe that

$$G(\Delta^{-1}(x)\mathbf{z}\Delta(x)) = G([z_{ij} \frac{x_j}{x_i}]) = \sum_{\mathbf{k}} g_{\mathbf{k}} \mathbf{z}^{\mathbf{k}} \prod_{i=1}^{m} x_i^{\sum_j k_{ji} - \sum_j k_{ij}}. \qquad (19)$$

Therefore, $G^*(\mathbf{z})$ is the coefficient of $G([z_{ij} \frac{x_j}{x_i}])$ at $x_1^0 x_2^0 \cdots x_m^0$ since $\sum_j k_{ji} - \sum_j k_{ij} = 0$ for matrices $\mathbf{k} \in \mathcal{F}$. We write it in shortly as $G^*(\mathbf{z}) = [x_1^0 \cdots x_m^0] g([z_{ij} \frac{x_j}{x_i}])$. The result follows from the Cauchy coefficient formula (cf. [21]). ∎

10

We consider the number of solutions to (4) and (5), which by Lemma 1 is asymptotically equivalent to the number of (cyclic) Markov types $|\mathcal{P}_n(m)|$ over the alphabet $\mathcal{A}$, and whose generating function is

$$F_m^*(z) = \sum_{n \geq 0} |\mathcal{F}_n(m)| z^n.$$

Then applying the above lemma with $z_{ij} = z x_i / x_j$ we conclude that

$$F_m^*(z) = \frac{1}{(1-z)^m} [x_1^0 x_2^0 \cdots x_m^0] \prod_{i \neq j} \left[1 - z \frac{x_i}{x_j} \right]^{-1}, \qquad (20)$$

since

$$F_m(z) = \sum_{\mathbf{k}} \mathbf{z}^{\mathbf{k}} = \prod_{ij} (1 - z_{ij})^{-1}.$$

Thus, by the Cauchy formula,

$$|\mathcal{F}_n(m)| = [z^n] F_m^*(z) = \frac{1}{2\pi i} \oint \frac{F_m^*(z)}{z^{n+1}} dz.$$

In the next section we evaluate asymptotically this expression to yield the following main result of this paper. Throughout we shall use the notation $f(n) \sim g(n)$ to mean $\lim_{n \to \infty} [f(n)/g(n)] = 1$.

Theorem 1 (i) *For fixed m and $n \to \infty$ the number of (cyclic) Markov types is*

$$|\mathcal{P}_n(m)| = d(m) \frac{n^{m^2 - m}}{(m^2 - m)!} + O(n^{m^2 - m - 1}) \qquad (21)$$

where $d(m)$ is a constant that also can be expressed by the following integral

$$d(m) = \frac{1}{(2\pi)^{m-1}} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{(m-1)-fold} \prod_{j=1}^{m-1} \frac{1}{1 + \phi_j^2} \cdot \prod_{k \neq \ell} \frac{1}{1 + (\phi_k - \phi_\ell)^2} \, d\phi_1 d\phi_2 \cdots d\phi_{m-1}. \qquad (22)$$

(ii) *When $m \to \infty$ we find that*

$$|\mathcal{P}_n(m)| \sim \frac{\sqrt{2} m^{3m/2} e^{m^2}}{m^{2m^2} 2^m \pi^{m/2}} \cdot n^{m^2 - m} \qquad (23)$$

provided that $m^4 = o(n)$.

Remark 1. It is easy to count the number of matrices $\mathbf{k}$ satisfying *only* equation (4), that is, $\sum_{ij} k_{ij} = n$. Indeed, it coincides with the number of integer solution of (4), which turns out to be the number of combinations with repetitions (the number of ways of selecting m^2 objects from n), that is,

$$\binom{n + m^2 - 1}{n} = \binom{n + m^2 - 1}{m^2 - 1} \sim \frac{n^{m^2 - 1}}{(m^2 - 1)!}$$

Thus the conservation law equation (5) decreases the above by the factor $\Theta(n^{m-1})$.

Table 1: Constants at n^{m^2-m} for fixed m and large m.

m	constant in (23)	constant in (21)
2	1.920140832 10^{-1}	2.500000000 10^{-1}
3	9.315659368 10^{-5}	1.157407407 10^{-5}
4	1.767043356 10^{-11}	2.174662186 10^{-11}
5	3.577891782 10^{-22}	4.400513659 10^{-22}

Remark 2. The evaluation of the integral (22) is quite cumbersome (see next section), but for small values of m we computed it to find that

$$|\mathcal{P}_n(2)| \quad \sim \quad \frac{1}{2}\frac{n^2}{2!} \tag{24}$$

$$|\mathcal{P}_n(3)| \quad \sim \quad \frac{1}{12}\frac{n^6}{6!} \tag{25}$$

$$|\mathcal{P}_n(4)| \quad \sim \quad \frac{1}{96}\frac{n^{12}}{12!} \tag{26}$$

$$|\mathcal{P}_n(5)| \quad \sim \quad \frac{37}{34560}\frac{n^{20}}{20!} \tag{27}$$

for large n. The coefficients of n^{m^2-m} are rational numbers since $F^*(z)$ is a rational generating function.

Remark 3. We now compare the coefficient at n^{m^2-m} for fixed m in (21) with its asymptotic counterpart in (23). They are shown in Table 1. Observe extremely small values of these constants even for relatively small m.

2.3 Counting Markov Types

Finally, we address the issue of Markov types over cyclic strings versus non-cyclic or linear strings. Recall that $\mathcal{Q}_n(m)$ denotes the set of Markov types with an arbitrary *initial condition* or simply over non-cyclic (linear) strings. We find a simple relation between $|\mathcal{Q}_n(m)|$ and $|\mathcal{P}_n(m)|$.

Let $a \in \mathcal{A}$ and define $\mathcal{P}_n(a, \mathcal{A})$ as the set of types over circular strings of length n that contain at least one occurrence of symbol a. Clearly,

$$|\mathcal{P}_n(a, \mathcal{A})| = |\mathcal{P}_n(m)| - |\mathcal{P}_n(m-1)|,$$

and therefore $|\mathcal{P}_n(a, \mathcal{A})| = |\mathcal{P}_n(m)|(1 - O(n^{-2m}))$ by Theorem 1.

In the same spirit, let $\mathcal{Q}_n(a)$ be the set of types over linear strings starting with symbol a, and $\mathcal{Q}_n(a, b)$ be the set of types over linear strings of length n that start with symbol a and

end with symbol b. Certainly, $\mathcal{Q}_n(a,a) = \mathcal{P}_{n-1}(a, \mathcal{A})$ and noticing that $\bigcup_{a \in \mathcal{A}} \mathcal{P}_n(a, \mathcal{A}) = \mathcal{P}_n(m)$, we conclude that

$$\mathcal{Q}_n(m) = \bigcup_{\substack{(a,b) \in \mathcal{A}^2 \\ a \neq b}} \mathcal{Q}_n(a,b) \cup \bigcup_{a \in \mathcal{A}} \mathcal{P}_{n-1}(a, \mathcal{A}) = \bigcup_{\substack{(a,b) \in \mathcal{A}^2 \\ a \neq b}} \mathcal{Q}_n(a,b) \ \cup \ \mathcal{P}_{n-1}(m). \tag{28}$$

Observe also that $\mathcal{Q}_n(a,b)$ are disjoint for $a \neq b$ and for every $a \neq b$ the set $\mathcal{Q}_n(a,b)$ is disjoint from $\mathcal{P}_n(a, \mathcal{A})$. Furthermore, cardinality of $\mathcal{Q}_n(a,b)$ is the same for all $a \neq b$. In summary, by (28) we conclude that

$$|\mathcal{Q}_n(m)| = |\mathcal{P}_{n-1}(m)| + (m^2 - m)|\mathcal{Q}_n(a,b)|. \tag{29}$$

We now estimate the cardinality of $\mathcal{Q}_n(a,b)$. We shall prove the following easy inequalities

$$|\mathcal{P}_{n-2}(a, \mathcal{A})| \ \leq \ |\mathcal{Q}_n(a,b)|, \tag{30}$$
$$|\mathcal{Q}_n(a,b)| \ \leq \ |\mathcal{P}_n(a, \mathcal{A})|. \tag{31}$$

Indeed, for (30) observe that if the count matrix $\mathbf{k} \in \mathcal{P}_{n-2}(a, \mathcal{A})$, then $\mathbf{k} + \mathbf{e}_{ab} \in \mathcal{Q}_n(a,b)$ where $\mathbf{e}_{ab}$ is a matrix with all 0's except at position (a,b) it contains a 1. This implies (30). For (31), we notice that if $\mathbf{k} \in \mathcal{Q}_n(a,b)$, then $\mathbf{k} + \mathbf{e}_{ba} \in \mathcal{P}_n(a, \mathcal{A})$ and the inequality follows.

Consequently, since $|\mathcal{P}_{n-2}(m)| \sim |\mathcal{P}_{n-1}(m)| = |\mathcal{P}_n(m)|(1 - O(n^{-m^2}))$, we have $|\mathcal{Q}_n(a,b)| = |\mathcal{P}_n(m)|(1 - O(n^{-2m}))$ which by (29) leads to our final conclusion.

Corollary 1 *Let $|\mathcal{P}_n(m)|$ denote the number of Markov types over cyclic strings as established in Theorem 1. The number of Markov types $|\mathcal{Q}_n(m)|$ with arbitrary initial conditions then satisfies*

$$|\mathcal{Q}_n(m)| = (m^2 - m + 1)|\mathcal{P}_n(m)|(1 - O(n^{-2m}))$$

where $|\mathcal{P}_n(m)|$ is given by (21).

3 Analysis and Proofs

In this section we prove Theorem 1. Our starting formula is (20) that we repeat below

$$F_m^*(z) = \frac{1}{(1-z)^m}[x_1^0 x_2^0 \cdots x_m^0] \prod_{i \neq j} \left[1 - z\frac{x_i}{x_j}\right]^{-1}. \tag{32}$$

3.1 Finite m

We first compute this explicitly for $m = 2, 3, 4, 5$ as summarized in Table 1.

For $m = 2$, we have

$$F_2^*(z) = \frac{1}{(1-z)^2}[x_1^0 x_2^0]\left[\frac{1}{1 - z\ x_1/x_2}\ \frac{1}{1 - z\ x_2/x_1}\right]. \tag{33}$$

Let us set $A = x_1/x_2$ so we need the coefficient of A^0 in $(1 - Az)^{-1}(1 - z/A)^{-1}$. Using a partial fractions expression in A, we have

$$\frac{1}{1-Az}\,\frac{1}{1-z/A} = \frac{1}{1-z^2}\left[\frac{1}{1-Az} + \frac{z}{A-z}\right].$$

For definiteness, we can assume that $|z| < |A| < |1/z|$ so that the coefficient of A^0 in $(1-Az)^{-1}$ is one and that in $z(A-z)^{-1}$ is zero. Hence, $F_2^*(z) = (1-z)^{-2}(1-z^2)^{-1} = (1+z)^{-1}(1-z)^{-3}$ and

$$
\begin{aligned}
|\mathcal{P}_n(2)| \sim |\mathcal{F}_n(2)| &= \frac{1}{2\pi i}\oint \frac{1}{z^{n+1}}\frac{1}{1+z}\,\frac{1}{(1-z)^3}dz \\
&= \frac{n^2}{4} + n + \frac{3}{4} + \frac{1}{8}[1+(-1)^n] \sim \frac{1}{2}\frac{n^2}{2!}, \qquad n\to\infty. \qquad (34)
\end{aligned}
$$

which agrees with (7) of Example 1.

For $m \geq 3$ we use recursive partial fractions expansions. When $m=3$ we set $x_1/x_2 = A$, $x_1/x_3 = B$ so that we wish to compute

$$[A^0 B^0]\left(\frac{1}{1-zA}\,\frac{1}{1-z/A}\,\frac{1}{1-Bz}\,\frac{1}{1-z/B}\,\frac{1}{1-Az/B}\,\frac{1}{1-Bz/A}\right). \qquad (35)$$

First we do a partial fractions expansion in the A variable, for fixed B and z. Thus the factor inside the parentheses in (35) becomes

$$
\begin{aligned}
&\ \frac{1}{1-zA}\,\frac{1}{1-z^2}\,\frac{1}{1-Bz}\,\frac{1}{1-z/B}\,\frac{1}{1-1/B}\,\frac{1}{1-Bz^2} \\
&+\ \frac{1}{1-z/A}\,\frac{1}{1-z^2}\,\frac{1}{1-Bz}\,\frac{1}{1-z/B}\,\frac{1}{1-z^2/B}\,\frac{1}{1-B} \\
&+\ \frac{1}{1-Az/B}\,\frac{1}{1-B}\,\frac{1}{1-z^2/B}\,\frac{1}{1-B/z}\,\frac{1}{1-z/B}\,\frac{1}{1-z^2} \\
&+\ \frac{1}{1-Bz/A}\,\frac{1}{1-Bz^2}\,\frac{1}{1-1/B}\,\frac{1}{1-Bz}\,\frac{1}{1-z/B}\,\frac{1}{1-z^2}. \qquad (36)
\end{aligned}
$$

The coefficient of A^0 in the first term in (36) is

$$\frac{1}{1-z^2}\,\frac{1}{1-Bz}\,\frac{1}{1-z/B}\,\frac{1}{1-1/B}\,\frac{1}{1-Bz^2}, \qquad (37)$$

and that in the third term is

$$\frac{1}{1-B}\,\frac{1}{1-z^2/B}\,\frac{1}{1-Bz}\,\frac{1}{1-z/B}\,\frac{1}{1-z^2}, \qquad (38)$$

while the coefficients of A^0 are zero in the second and fourth terms. Combining (37) and (38) we must now compute

$$[B^0]\left(\frac{1+z^2}{1-z^2}\,\frac{1}{1-Bz}\,\frac{1}{1-z/B}\,\frac{1}{1-Bz^2}\,\frac{1}{1-z^2/B}\right). \qquad (39)$$

Now expanding (39) by a partial fractions expansion in B leads to

$$\frac{1+z^2}{1-z^2}[B^0]\left(\frac{1}{1-Bz}\frac{1}{1-z^2}\frac{1}{1-z}\frac{1}{1-z^2}+\frac{1}{1-z/B}\frac{1}{1-z^2}\frac{1}{1-z^3}\frac{1}{1-z}\right.$$

$$+\left.\frac{1}{1-1/z}\frac{1}{1-z^3}\frac{1}{1-Bz^2}\frac{1}{1-z^4}+\frac{1}{1-z^3}\frac{1}{1-1/z}\frac{1}{1-z^4}\frac{1}{1-z^2/B}\right)$$

$$=\frac{1+z^2}{1-z^2}\left[\frac{1}{1-z^2}\frac{1}{1-z}\frac{1}{1-z^3}+\frac{-z}{(1-z)}\frac{1}{1-z^3}\frac{1}{1-z^4}\right]=\frac{1-z+z^2}{(1-z)^4(1+z)^2(1+z+z^2)}.$$

Hence,

$$F_3^*(z)=\frac{1-z+z^2}{(1-z)^7(1+z)^2(1+z+z^2)}.$$

For $z\to1$, $F_3^*(z)\sim\frac{1}{12}(1-z)^{-7}$ so that

$$|\mathcal{P}_n(3)|\sim\frac{1}{12}\frac{n^6}{6!},\qquad n\to\infty.\tag{40}$$

Using similar recursive partial fractions expansions, with the help of the symbolic computation program MAPLE, we find that for $m=4$ and $m=5$

$$F_4^*(z)=\frac{z^8-2z^7+3z^6+2z^5-2z^4+2z^3+3z^2-2z+1}{(1-z)^{13}(1+z)^5(1+z^2)(1+z+z^2)^2}\tag{41}$$

and

$$F_5^*(z)=\frac{Q(z)}{(1-z)^{21}(1+z)^8(1+z^2)^2(1+z+z^2)^4(1+z+z^2+z^3+z^4)},\tag{42}$$

where

$$\begin{aligned}Q(z)&=z^{20}-3z^{19}+7z^{18}+3z^{17}+2z^{16}+17z^{15}+35z^{14}+29z^{13}+45z^{12}+50z^{11}\\&+72z^{10}+50z^9+45z^8+29z^7+35z^6+17z^5+2z^4+3z^3+7z^2-3z+1.\end{aligned}$$

These results show that it is unlikely that a simple formula can be found for $F_m^*(z)$ for general m.

By expanding (41) and (42) near $z=1$ we conclude that as $n\to\infty$

$$|\mathcal{P}_n(4)|\sim\frac{1}{96}\frac{n^{12}}{12!},\qquad|\mathcal{P}_n(5)|\sim\frac{37}{34560}\frac{n^{20}}{20!}.\tag{43}$$

It is easy to inductively show that at $z=1$, $F_m^*(z)$ has a pole of order m^2-m+1 and the other singularities are poles at the roots of unity that are of order $<m^2-m+1$. These poles and their orders are given in Table 2.

Thus, for $n\to\infty$, we have

$$|\mathcal{P}_n(m)|\sim d(m)\frac{n^{m^2-m}}{(m^2-m)!},\tag{44}$$

where

$$d(m)=\lim[(1-z)^{m^2-m+1}F_m^*(z)]$$

15

Table 2: Poles and their orders for various m.

$m\backslash$ root	1	–1	$e^{\pm 2\pi i/3}$	$\pm i$	$e^{\pm 2\pi i/5}$	$e^{\pm 4\pi i/5}$
2	3	1	–	–	–	–
3	7	2	1	–	–	–
4	13	5	2	1	–	–
5	21	8	4	2	1	1

as $z \to 1$. However, there seems to be no simple formula for the sequence of constants $d(m)$. We proceed to characterize $d(m)$ as an $(m-1)$ fold integral.

First consider the simple case $m = 2$. Setting $A = e^{i\Phi}$ and using a Cauchy integral, we have

$$[A^0] \; \frac{1}{1 - z/A} \; \frac{1}{1 - Az} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{d\Phi}{1 - 2z\cos\Phi + z^2}.$$

Now set $z = 1 - \delta$ and expand the integral for $z \to 1$. The major contribution will come from where $\delta \approx 0$ and scaling $\Phi = \delta\phi$ and using the Taylor expansion $1 - 2(1 - \delta)\cos(\delta\phi) + (1 - \delta)^2 = \delta^2[1 + \phi^2] + O(\delta^3)$, we find that

$$F_2^*(z) \sim \frac{1}{\delta^2} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\delta}{\delta^2[1 + \phi^2]} d\phi = \frac{1}{2} \frac{1}{\delta^3}, \qquad \delta \to 0.$$

When $m = 3$, we use (3.4) and the Cauchy integral formula with $A = e^{i\Phi}$ and $B = e^{i\Psi}$ to get

$$\frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{1}{1 - 2z\cos\Phi + z^2} \cdot \frac{1}{1 - 2z\cos\Psi + z^2} \cdot \frac{1}{1 - 2z\cos(\Phi - \Psi) + z^2} \; d\Phi d\Psi.$$

Again expanding the above for $z = 1 - \delta \to 1$ and $\Phi = \delta\phi = O(\delta)$, $\Psi = \delta\psi = O(\delta)$, we obtain the leading order approximation

$$\frac{1}{\delta^4} \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{1 + \phi^2} \frac{1}{1 + \psi^2} \frac{1}{1 + (\phi - \psi)^2} \; d\phi d\psi = \frac{1}{\delta^4} \cdot \frac{1}{12}.$$

Thus as $z \to 1$, $F_3^*(z) \sim \frac{1}{12}\delta^{-7} = \frac{1}{12}(1 - z)^{-7}$ which follows also from the exact generating function.

For general m a completely analogous calculation shows that as $\delta = 1 - z \to 0$, $F_m^*(z) \sim \delta^{m-m^2-1}d(m)$ where

$$d(m) = \frac{1}{(2\pi)^{m-1}} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{(m-1)-fold} \prod_{j=1}^{m-1} \frac{1}{1 + \phi_j^2} \cdot \prod_{k \neq \ell} \frac{1}{1 + (\phi_k - \phi_\ell)^2} \; d\phi_1 d\phi_2 \cdots d\phi_{m-1}. \quad (45)$$

16

The second product in the above is over all distinct pairs (k, ℓ), so that this may also be written as

$$\prod_{\ell=1}^{m-2} \prod_{k=\ell+1}^{m-1} \frac{1}{1 + (\phi_k - \phi_\ell)^2}. \tag{46}$$

This completes the proof of part (i) of Theorem 1.

3.2 Large m

We now use the saddle point method ([21] (Chap. 8.4) and [12] Chap. VIII) to prove part (ii) of Theorem 1. Since

$$\sum_{k_{ij}} \left(z \frac{z_i}{z_j} \right)^{k_{ij}} = \left(1 - z \frac{z_i}{z_j} \right)^{-1}$$

and setting $z_i = e^{i\theta_i}$ we find that

$$F_m^*(z) = \frac{1}{(2\mathbf{i}\pi)^m} \oint \cdots \oint \prod_{ij} \left(1 - z \frac{z_i}{z_j} \right)^{-1} \frac{dz_1}{z_1} \cdots \frac{dz_m}{z_m} \tag{47}$$

$$= (2\pi)^{-m} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \prod_{ij} (1 - z \exp(i(\theta_i - \theta_j)))^{-1} d\theta_1 \cdots d\theta_m . \tag{48}$$

By noticing that the expression $\prod_{ij}(1 - z \exp(i(\theta_i - \theta_j)))^{-1}$ does not change when the θ_i are all incremented by the same value, one can integrate over θ_1 to obtain

$$F_m^*(z) = (2\pi)^{-m+1} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \prod_i (1 - z \exp(\mathbf{i}\theta_i))^{-1} (1 - z \exp(-\mathbf{i}\theta_i))^{-1}$$

$$\times \frac{1}{1-z} \prod_{i>1, j>1} (1 - z \exp(\mathbf{i}(\theta_i - \theta_j)))^{-1} d\theta_2 \cdots d\theta_m. \tag{49}$$

Let now

$$L(z, \theta_2, \ldots, \theta_m) = \log(1 - z) + \sum_i \log(1 - z \exp(\mathbf{i}\theta_i))(1 - z \exp(-\mathbf{i}\theta_i))$$

$$+ \sum_{i>1, j>1} \log(1 - z \exp(\mathbf{i}(\theta_i - \theta_j))).$$

An alternative form of the above is

$$L(z, \theta_2, \ldots, \theta_m) = \log(1 - z) + \sum_{i=2}^{m} \log(1 - 2z \cos \theta_i + z^2)$$

$$+ \frac{1}{2} \sum_{i=2}^{m} \sum_{j=2}^{m} \log(1 - 2z \cos(\theta_i - \theta_j) + z^2).$$

Notice that $L(z, 0, \ldots, 0) = m^2 \log(1 - z)$. Hence

$$|\mathcal{F}_n(m)| = \frac{1}{\mathbf{i}(2\pi)^m} \oint \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \exp(-L(z, \theta_2, \ldots, \theta_m)) \frac{dz}{z^{n+1}} d\theta_2 \cdots d\theta_m. \tag{50}$$

17

In order to find the asymptotics of this integral we use the *multidimensional saddle point method*. The quantity $L(z, \theta_2, \ldots, \theta_m) + n \log z$ attains its minimum value at $(\theta_2, \ldots, \theta_m) = (0, \ldots, 0)$ and $z = z_n = \frac{n}{m^2+n}$. The minimum value is therefore

$$m^2 \log(1 - z_n) + n \log z_n = m^2 \log(m^2/(m^2 + n)) + n \log(n/(m^2 + n))$$

or

$$m^2 \log(m^2) + n \log(n) - (m^2 + n) \log(m^2 + n).$$

Then

$$m^2 \log(1 - z_n) + n \log z_n = m^2 \log m^2 - m^2 \log n - m^2 + O(m^4/n)$$

provided that $m^4 = o(n)$.

It turns out that at $(z, \theta_2, \ldots, \theta_m) = (z_n, 0, \ldots, 0)$ we have:

$$\frac{\partial^2}{\partial z^2} L(z, \theta_2, \ldots, \theta_m) = -\frac{m^2}{(1 - z_n)^2}$$

$$\forall i: \frac{\partial^2}{\partial z \partial \theta_i} L(z, \theta_2, \ldots, \theta_m) = 0$$

$$\forall i: \frac{\partial^2}{\partial \theta_i^2} L(z, \theta_2, \ldots, \theta_m) = \frac{2(m - 1)z_n}{(1 - z_n)^2}$$

$$\forall i \neq j: \frac{\partial^2}{\partial \theta_i \partial \theta_j} L(z, \theta_2, \ldots, \theta_m) = -\frac{2z_n}{(1 - z_n)^2}, \quad m \geq 3.$$

In other words, the second derivative matrix $\mathbf{Q}_2$ of $L(z, \theta_2, \ldots, \theta_m) + n \log z$ at $(z, \theta_2, \ldots, \theta_m) = (z_n, 0, \ldots, 0)$ is

$$\mathbf{Q}_2 = \left(-\frac{m^2}{(1 - z_n)^2} - \frac{n}{(z_n)^2}\right) \mathbf{u}_z \otimes \mathbf{u}_z + \frac{2mz_n}{(1 - z_n)^2} \mathbf{I}_\theta - \frac{2z_n}{(1 - z_n)^2} \mathbf{u}_\theta \otimes \mathbf{u}_\theta$$

where $\mathbf{u}_z = (1, 0, \ldots, 0)$,

$$\mathbf{u}_\theta = \frac{1}{\sqrt{m - 1}}(0, 1, \ldots, 1),$$

and

$$\mathbf{I}_\theta = \mathbf{I} - \mathbf{u}_z \otimes \mathbf{u}_z,$$

i.e., the identity restricted on θ components. In the above $\otimes$ is the tensor product (in our case, it is a product of two vectors resulting in a matrix). For example,

$$\mathbf{u}_\theta \otimes \mathbf{u}_\theta = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 1 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 1 & \ldots & 1 \end{pmatrix}.$$

An application of the saddle point method yields

$$|\mathcal{F}_n(m)| \sim \frac{1}{(2\pi)^{m/2} z_n \sqrt{\det(\mathbf{Q}_2)}} \exp(-m^2 \log(1 - z_n) - n \log z_n)$$

where $\det(\cdot)$ denotes the determinant. Since

$$|\det(\mathbf{Q}_2)| = \left(\frac{m^2}{(1-z_n)^2} + \frac{n}{(z_n)^2}\right)\left(\frac{z_n}{(1-z_n)^2}\right)^{m-1} 2^{m-1}m^{m-2} \sim n^{2m}m^{-3m}2^{m-1},$$

we find that for $m^4 = o(n)$

$$|\mathcal{F}_n(m)| \sim |\mathcal{P}_n(m)| \sim \left(m^{-2m^2+3m/2}e^{m^2}2^{-m}\pi^{-m/2}\sqrt{2}\right)n^{m^2-m},$$

and this completes the proof. The condition $m^4 = o(n)$ is needed since we used the approximation for $m^2\log(1-z_n)$ below (50).

Acknowledgment

References

[1] K. Atteson, The Asymptotic Redundancy of Bayes Rules for Markov Chains, *IEEE Trans. on Information Theory*, 45, 2104-2109, 1999.

[2] A. Barron, J. Rissanen, and B. Yu, The Minimum Description Length Principle in Coding and Modeling, *IEEE Trans. Information Theory*, 44, 2743-2760, 1998.

[3] P. Billingsley, Statistical Methods in Markov Chains, *Ann. Math. Statistics*, 32, 12-40, 1961.

[4] T. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, 1991.

[5] I. Csziszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York, 1981.

[6] I. Csziszár, The Method of Types, *IEEE Trans. Information Theory*, 44, 2505-2523, 1998.

[7] L. D. Davisson, Universal Noiseless Coding, *IEEE Trans. Information Theory*, 19, 783-795, 1973.

[8] L. D. Davisson, Minimax Noiseless Universal coding for Markov Sources, *IEEE Trans. Information Theory*, 29, 211 - 215, 1983.

[9] M. Drmota and W. Szpankowski, Precise Minimax Redundancy and Regret, *IEEE Trans. Information Theory*, 50, 2686-2707, 2004.

[10] M. Feder, N. Merhav, and M. Gutman, Universal Prediction of Individual Sequences, *IEEE Trans. Information Theory*, 38, 1258–1270, 1992.

[11] J. Fill and V. Lyzinski, Counting Markov Types: Comments, Corrections, Extensions, and Problems, Private Communication, 2010.

[12] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2008.

[13] L. Goodman, Exact Probabilities and Asymptotic Relationships for Some Statistics from m-th Order Markov Chains, *Annals of Mathematical Statistics*, 29, 476-490, 1958.

[14] P. Jacquet and W. Szpankowski, Markov Types and Minimax Redundancy for Markov Sources, *IEEE Trans. Information Theory*, 50, 1393-1402, 2004.

[15] A. Martín, G. Seroussi, and M. J. Weinberger, Type classes of tree models, *Proc. ISIT 2007*, Nice, France, 2007.

[16] J. Rissanen, Fisher Information and Stochastic Complexity, *IEEE Trans. Information Theory*, 42, 40-47, 1996.

[17] G. Seroussi, On Universal Types, *IEEE Trans. Information Theory*, 52, 171-189, 2006.

[18] P. Shields, Universal Redundancy Rates Do Not Exist, *IEEE Trans. Information Theory*, 39, 520-524, 1993.

[19] R. Stanley, *Enumerative Combinatorics*, Vol. II, Cambridge University Press, Cambridge, 1999.

[20] W. Szpankowski, On Asymptotics of Certain Recurrences Arising in Universal Coding, *Problems of Information Transmission*, 34, 55-61, 1998.

[21] W. Szpankowski, *Average Case Analysis of Algorithms on Sequences*, Wiley, New York, 2001.

[22] Q. Xie, A. Barron, Asymptotic Minimax Regret for Data Compression, Gambling, and Prediction, *IEEE Trans. Information Theory*, 46, 431-445, 2000.

[23] M. J. Weinberger, N. Merhav, and M. Feder, Optimal sequential probability assignment for individual sequences, *IEEE Trans. Inf. Theory*, 40, 384-396, 1994.

[24] P. Whittle, Some Distribution and Moment Formulæ for Markov Chain, *J. Roy. Stat. Soc.*, Ser. B., 17, 235-242, 1955.

Philippe Jacquet graduated form Ecole Polytechnique, Paris, France in 1981, and from Ecole des Mines in 1984. He received his PhD degree from Paris Sud University in 1989. Since 1998, he has been a research director in Inria, a major public research lab in Computer Science in France. He has been a major contributor to the Internet OLSR protocol for mobile networks. His research interests involve information theory, probability theory, quantum telecommunication, protocol design, performance evaluation and optimization, and the analysis of algorithms. Since 2012 he is with Alcatel-Lucent Bell Labs as head of the department of mathematics of dynamic networks and information.

Charles Knessl is currently Professor of Applied Mathematics at the University of Illinois at Chicago. He is the past recipient of NSF Graduate and Postdoctoral Research Fellowships, a Sloan Fellowship and an NSF Presidential Young Investigator Award. His research interests include asymptotic and singular perturbation methods, queuing theory, risk theory, problems of convection-diffusion, the analysis of algorithms, financial mathematics and applications of special functions. He is the author or co-author on approximately 170 publications.

Wojciech Szpankowski is Saul Rosen Professor of Computer Science and (by courtesy) Electrical and Computer Engineering at Purdue University where he teaches and conducts research in analysis of algorithms, information theory, bioinformatics, analytic combinatorics, random structures, and stability problems of distributed systems. He received his M.S. and Ph.D. degrees in Electrical and Computer Engineering from Gdansk University of Technology. He held several Visiting Professor/Scholar positions, including McGill University, INRIA, France, Stanford, Hewlett-Packard Labs, Universite de Versailles, University of Canterbury, New Zealand, Ecole Polytechnique, France, and the Newton Institute, Cambridge, UK. He is a Fellow of IEEE, and the Erskine Fellow. In 2010 he received the Humboldt Research Award. In 2001 he published the book "Average Case Analysis of Algorithms on Sequences", John Wiley & Sons, 2001. He has been a guest editor and an editor of technical journals, including THEORETICAL COMPUTER SCIENCE, the ACM TRANSACTION ON ALGORITHMS, the IEEE TRANSACTIONS ON INFORMATION THEORY, FOUNDATION AND TRENDS IN COMMUNICATIONS AND INFORMATION THEORY, COMBINATORICS, PROBABILITY, AND COMPUTING, and ALGORITHMICA. In 2008 he launched the interdisciplinary Institute for Science of Information, and in 2010 he became the Director of the newly established NSF Science and Technology Center for Science of Information.